

# Learning high-level spectral-spatial features for hyperspectral image classification with insufficient labeled samples

Douglas Omwenga Nyabuga, Godfrey Nyariki

Department of Information Technology, School of Computing, Informatics and Media Studies, Mount Kigali University, Kigali, Rwanda

## Article Info

### Article history:

Received Apr 24, 2024

Revised Nov 7, 2024

Accepted Nov 14, 2024

### Keywords:

Classification

DenseNet

Hyperspectral image

Principal component analysis

Spectral-spatial

## ABSTRACT

Hyperspectral image (HSI) classification research is a hot area, with a mass of new methods being developed to improve performance for specific applications that use spatial and spectral image material. However, the main obstacle for scientists is determining how to identify HSIs effectively. These obstacles include an increased presence of redundant spectral information, high dimensionality in observed data, and limited spatial features in a classification model. To this end, we, therefore, proposed a novel approach for learning high-level spectral-spatial features for HSI classification with insufficient labeled samples. First, we implemented the principal component analysis (PCA) technique to reduce the high dimensionalities experienced. Second, a fusion of 2D and 3D convolutions and DenseNet, a transfer learning network for feature learning of both spatial-spectral pixels. The achieved experimental results are comparatively satisfactory to contrasted approaches on the widely used HSI images, i.e., the University of Pavia and Indian Pines, with an overall classification accuracy of 97.80% and 97.60%, respectively.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



## Corresponding Author:

Douglas Omwenga Nyabuga

Department of Information Technology, School of Computing, Informatics and Media Studies

Mount Kigali University

P.O Box 5628, Kigali, Rwanda

Email: dnyabuga@mkurwanda.ac.rw

## 1. INTRODUCTION

Hyperspectral imagery (HSI) measures reflectance values of the electromagnetic spectra in over a hundred spectral bands to every spatial region in the image. While these valuable spectral details improve the capacity to distinguish objects, HSI analysis needs more complex algorithms because of the high dimension of the pixels, high nonlinearity, and the small-sample problem of HSI data [1], [2]. Therefore, many researchers, for example in [3], [4] have explored these HSI dimensionality reduction techniques. However, there is an increase in variation in spatial dimension.

Hyperspectral sensors produce massive volumes of data, resulting in a large volume of bands in the data, making real-time parameters difficult and laborious to achieve. As a result, it is a desirable strategy to reduce the data size before starting high-level processing. Therefore, in recent years, a dimensionality reduction stage has become a significant part of machine learning (ML). Further, the research in [5], [6] mainly carried out dimensionality reduction using the principal components transformation, which chose and preserved the most relevant data for classification. As a result, classifiers create efficient models at minimal computational cost and enhance pixel classification accuracy in HSIs. Nevertheless, there is a challenge in how to make full use of the spatial-spectral features contained in HSI to improve HSI classification.

For this reason, the principal component analysis (PCA) has recently been used as a multivariable approach for dimensionality reduction [5], [7], [8]. It is the most adopted methodology in remote sensing applications, mainly those applying HSIs. The adjacent bands are highly correlated in this type of image, hence gaining little additional information. PCA minimizes the amount of data by reducing dependencies between the various bands. An eigenvector decomposition of the original data's covariance matrix is computed to achieve this [6]. However, PCA only seeks the best orthogonal vectors, omitting crucial features essential for HSI classification.

Since its conception, the classification of HSIs has drawn widespread attention and spawned a plethora of approaches aimed at allocating a pixel (or a spectrum) to one of a set of classes [9], [10]. Several approaches in the literature have focused on investigating the importance of HSI data spectral signatures in classification, using only the spectrum of a pixel to establish its class membership. However, two fundamental difficulties for such pixel-wise techniques benefit from relative conceptual simplicity and implementation ease: i) the limited training set compared to the high-dimensional spectra and ii) the spectral variations. The first issue, which has been extensively studied in light of the well-known Hughes phenomenon [11], causes problems in two ways. First, due to the limited number of labeled samples, the sample covariance matrix is likely to be singular, resulting in ill-posed difficulties for several classification algorithms. Second, high-dimensional spectra require numerous free parameters for computation in a parametric approach, which is inclined to overfit and consequently decreases the generalization capacity of classifiers. Regarding spectral variation caused by several factors such as incident light, atmospheric effects, undesirable shade and shadow, natural spectrum fluctuation, and instrument noises [9], [12], two serious challenges might make categorization difficult.

On the one hand, substantial intra-class spectra variability makes it challenging to identify a specific class. Besides, low inter-class spectral variation makes distinguishing distinct classes difficult. These issues make HSI classification complex, resulting in poor classification results when using pixel-wise approaches. Because HSIs are naturally 3-D and visual, spatial reliance, analogous to spectral behavior, is a natural complement to spectra. As a result, the inclusion of spatial dependency has the potential to improve pixel-wise classification. The use of spatial information in HSI classification dates back over a decade, and some successful research has demonstrated its ability to enhance classification performance [13]. Since then, there has been a significant increase in interest in spectral-spatial classification. Scholars have utilized multilayer strategies to solve this challenge. However, these methods only analyze 2D and 3D spectral-spatial properties independently and take a long time to complete.

To this end, we present a spectral-spatial HSI classification approach based on DenseNet [14], which we employ as a unique strategy for HSI dataset classification. Furthermore, we incorporated the DenseNet framework for its standardization technique, which has numerous advantages: i) the reusability of the information (feature) covered in HSIs, ii) the concatenation of different paths (which reduces the number of parameters), iii) aided to overcoming problems such as overfitting and the vanishing gradient when few training samples are available, and iv) strengthen feature propagation. The summary of the main contributions of our proposed approach is:

- We implemented the PCA technique on the HSI for dimensionality reduction due to the high dimensions involved in the HSIs.
- Our proposed approach introduced the DenseNet-based mechanism to model the semantic interdependencies in spatial and channel dimensions to improve feature representation for classification ability.
- The achieved results demonstrate that our suggested approach can be trained end-to-end and estimated as the state-of-the-art for both datasets concurrently.

The remaining parts of this study are examined as follows: section 2 discusses related works. Section 3 discusses our proposed approach. Section 4 presents the experiments, i.e., the dataset and performance analysis metrics, and compares baseline methods, results, and discussions. Finally, section 5 concludes our study by giving a summary of the content and future recommendations.

## 2. RELATED WORK

Lately, many researchers have deployed convolutional neural networks (CNNs) in HSI categorization. They are better suited for HSI analysis and feature extraction [15], [16]. Though these approaches try to optimize the use of both spectral-spatial features, they usually divide the joint spatial-spectral features into two independent learning components, ignoring the relation underlying the spectral-spatial features.

The research in [17], [18] described the few-shot learning technique, in which the model efficiently discriminated categories in a newly acquired data set using only a small number of labeled samples.

However, all of these approaches rely on artificially calculated measurement distances, which may only partially apply to the features retrieved by the neural network when categorizing them. As a result, Jia *et al.* [19] proposed an effective transfer learning strategy to address inadequate training HSI samples. Even though this strategy has achieved significant advances in HSI classification, it performs poorly when only a few labeled samples are available. This has resulted in a significant issue for deep learning (DL) models, as addressed in this work. The research in [20], [21] state that optical remote sensing collects radiation reflected and emitted from the surfaces under study, focusing on the region of the electromagnetic spectrum with wavelengths spanning from visible to near-infrared to thermal infrared. With a HSI that captures a variety of precisely calibrated tiny spectral bands of the visible and infrared spectrums. The enormous amount of spectral data provides important land-cover information that helps precisely classify surface land use and land cover. Nevertheless, labor and time-intensive procedures are required to extract tagged training data from HSI. Consequently, based on active learning, a classifier design that uses the fewest labeled examples as practical for classification was proposed [22]. Remote sensing imagery (RSI) objects and features frequently have unclear backgrounds and cannot yield helpful information. Identifying the RSI is more difficult because of the notable intraclass variances.

Guo *et al.* [23] presented a multi-view-feature-learning network to address this problem and gather three specific domain features for the scene categorization challenge. On the other hand, Pundir and Akshay [24] introduced model-agnostic meta-learning and the ensemble of prototype networks to overcome the problems associated with standard deep-learning networks. This technique tackled the RSI categorization problem by applying meta-learning. Multiresolution categorization of panchromatic and multispectral pictures is a popular area of research. The main challenge in this field is assessing data and extracting characteristics to improve classification accuracy properly.

An adaptive hybrid fusion network that incorporates both data fusion and feature fusion was presented in [25] to classify multiresolution RSI. However, these methods depend on a large number of labeled training samples to obtain an excellent classification performance. To solve the classification problem, Sathyanarayana and Singh [26] designed a multilayer feedforward artificial neural network and used a histogram technique to extract the pixel density distribution and normalization to make the result independent from the physical properties of the image. Liu *et al.* [27] suggested a technique that combines a low-resolution HSI with a high-resolution (HR) multi-spectral image (MSI) to extract deep multiscale properties from an HSI scene. Training data was optional for this strategy. This research aims to reliably classify HSI into a class or category regardless of source, resolution, or size by developing a spectral-spatial approach based on the DenseNet network. Thus, it will expedite the process and enhance speed.

### 3. METHOD

This section covers the details of our proposed approach for spectral-spatial HSI classification. Figure 1 demonstrates the general framework of the proposed method. The input of our model is hyperspectral data with the  $b$  – spectral band and size of  $m \times n$ . Thus, we consider it as the matrix of order  $m \times n \times b$ . The PCA is applied to the HSI for data dimensionality reduction. The 3D spectral-spatial pixels are convoluted or concatenated. We introduced a DenseNet layer before the full convolutional (FC) layers for 3D spectral-spatial feature learning.

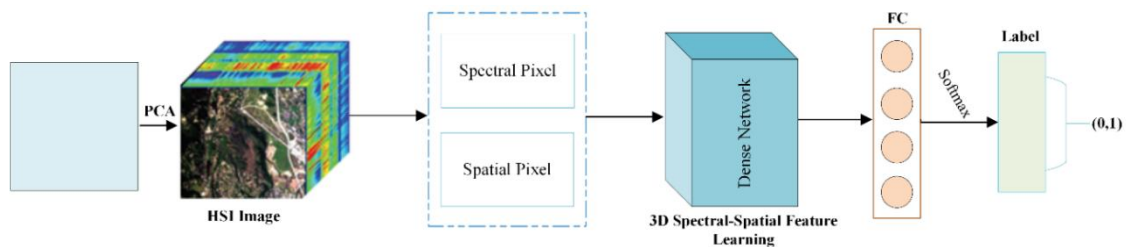


Figure 1. Proposed scheme for spatial-spectral classification of HSI

#### 3.1. Low-space projections

For the HSI dimensionality reduction through PCA takes the mathematical formulation of (1).

$$I = P^{m \times n \times b} \quad (1)$$

Let  $X \in \mathbb{R}^{r \times c \times n}$  represent the raw HSI, with  $r, c$ , and  $n$  indicating the row, column, and band numbers, respectively. Only the first  $p$  principal components are reserved when the PCA is applied to the raw image to minimize the convolution phase's computational costs. The dimension-reduced image is represented by  $X_p \in \mathbb{R}^{r \times c \times p}$ . A neighbor zone with the dimensions  $w \times w \times p$  is extracted around each pixel. Suppose we take  $m$  training samples; then  $X_{train} \in \mathbb{R}^{m \times w \times w \times p}$  denotes the training set.

### 3.2. Spectral pixel extraction

HSIs have a lot of spectrum information and spectral resolution; hence, their classification approach is based on spectral features. Each pixel can extract 1D spectral vectors to classify objects. We deployed the 1D-CNN to extract the spectral features from HSIs and categorize them, as demonstrated in Figure 2.

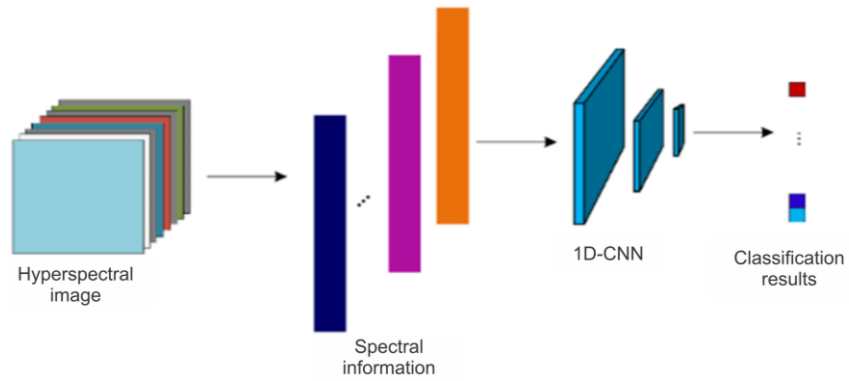


Figure 2. The spectral pixel extraction

The convolutional layer is introduced first. The value of neuron  $v_{i,j}^x$  at a position  $x$  of the  $j$ th feature map in the  $i$ th layer is defined as (2).

$$v_{i,j}^x = g \left( \sum_m \sum_{p=0}^{P_i-1} w_{i,j,m}^p v_{(i-1)m}^{x+p} + b_{i,j} \right) \quad (2)$$

Where  $m$  indexes the feature map in the previous layer ( $(i-1)$ th layer) connected to the current feature map,  $w_{i,j,m}^p$  denotes the weight of position  $p$  connected to the  $m$ th feature map,  $P_i$  denotes the width of the kernel toward the spectral dimension, and  $b_{i,j}$  denotes the bias of the  $j$ th feature map in the  $i$ th layer.

### 3.3. Spatial pixel extraction

Spatial pixels, i.e., context information, which forms part of the HSI images, are used to classify the HSIs. The spatial pixels extracted from the contiguous of the pixel are utilized instead of the spectral features derived from a specific pixel. Due to the wide range of hyperspectral data, the standard strategy for obtaining 2D (spatial) features is to condense the dataset first, use a two-dimensional CNN to obtain more detailed spatial features, and classify using spatial details. Figure 3 depicts the specific procedure. To execute a convolution operation on 2D data in the 2D convolution procedure, we applied a 2D convolution kernel. In the 2D convolution operation, input data is convolved with 2D kernels, and the process can be formulated as (3).

$$v_{i,j}^{x,y} = g \left( \sum_m \sum_{h=0}^{H_{j-1}} \sum_{w=0}^{W_{j-1}} k_{i,j,m}^{hw} v_{(i-1)m}^{(x+h)(y+w)} + b_{i,j} \right) \quad (3)$$

### 3.4. Spectral-spatial pixel extraction

Sole spectral information is usually employed in traditional HSI for classification purposes. Because of the influence of the natural atmosphere, identical land features will display different spectral curves. The so-known alike-object hetero-spectrum and unrelated-object but with similar spectrum phenomena can cause different ground objects to have the same spectral curve. For example, specific pixels (elements) interconnected on earth are designated as parking lots; therefore, pixels with spectral features that look highly similar to metal spectral features are most likely to represent vehicles. If a pixel has many grass pixels surrounding it, the pixels in the middle are most likely grass. Hyperspectral data has a 3D structure that

includes 1D (spectral features) and 2D (spatial features). A 3D-CNN may extract both spectral and spatial information at the same time. Figure 4 depicts the procedure for extracting the spectral-spatial features.

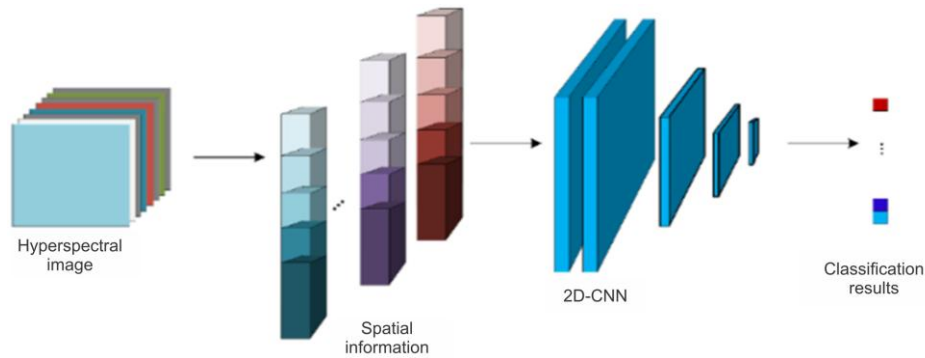


Figure 3. The spatial pixel extraction

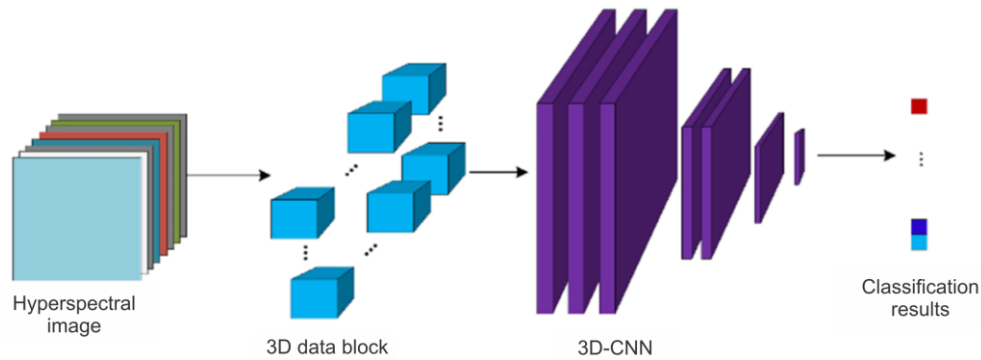


Figure 4. The spectral-spatial pixel extraction

The concatenation of the 2D and 3D convolutions of the HSI, it's defined through in (4).

$$v_{lij}^{xyz} = g \left( \sum_{h=0}^{H_l-1} \sum_{w=0}^{W_l-1} \sum_{d=0}^{D_l-1} k_{lj}^{hwd} v_{(l-1)i}^{(x+h)(y+w)(z+d)} + b_{lj} \right) \quad (4)$$

where  $D_l$  denotes the spectral depth of the 3-D kernel,  $i$  represents the number of feature cubes in the prior layer,  $j$  denotes the number of kernels in this layer.  $v_{lij}^{xyz}$  denotes the output at position  $(x, y, z)$ , which is computed by convolving the  $i$ th feature cube of the previous layer with the  $j$ th kernel of the  $l$ th layer, and  $k_{lj}^{hwd}$  is the  $(h, w, d)$ th value of the kernel connected to the  $i$ th feature cube in the preceding layer. As such, the output data of the  $l$ th convolution layer comprises  $i \times j$  3-D feature cubes.

In our model, we found out the expected labels via the FC layers and a SoftMax layer. Then, the loss function of the entire network is formulated using (5).

$$L = \frac{1}{m} \sum_{i=1}^m [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (5)$$

where  $y_i$  and  $\hat{y}_i$  denote the label and predicted of the  $i$ th data, respectively.  $m$  represents the number of training samples.

Our approach adopted the rectified linear unit (ReLU) function  $g(x) = \max(0, x)$ . ReLU activation guarantees the convolutional feature extractors (FE) nonlinearity and aids the quicker training of the network. In addition, we deployed the mini-batch stochastic gradient descent (SGD) approach to optimize the network effectively. To train our dataset in the experiments, we set the training epochs to 100, the learning rate (lr) to 0.001, and the patch size to 25. In addition, all simulations were executed on a MacBook Pro laptop with an Intel i7-5820K 3.30 GHz processor, 8 GB of RAM, and an NVIDIA GTX1080 graphics card and GPU

(Colab with 25 GB RAM), all running on Python 3.9. Figure 5, thus depicts the summary of our proposed model parameters.

Layer (type)	Output Shape	Param #	Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(25, 25, 30, 1)	0	conv4_4_0 (Conv2D)	(22, 22, 32)	4128
conv1 (Conv3D)	(25, 25, 30, 8)	512	conv4_4_1 (Conv2D)	(22, 22, 32)	9248
conv1_1 (Conv3D)	(24, 24, 29, 1)	144	conv4_2_1 (Conv2D)	(22, 22, 32)	9248
conv2_1_0 (Conv3D)	(24, 24, 29, 4)	68	conv4_3_1 (Conv2D)	(22, 22, 32)	9248
conv2_2_0 (Conv3D)	(24, 24, 29, 4)	68	conv4_4_1 (Conv2D)	(22, 22, 32)	9248
conv2_3_0 (Conv3D)	(24, 24, 29, 4)	68	conv5_1_0 (Conv2D)	(11, 11, 32)	4128
conv2_4_0 (Conv3D)	(24, 24, 29, 4)	68	conv5_2_0 (Conv2D)	(11, 11, 32)	4128
conv2_1_1 (Conv3D)	(24, 24, 29, 4)	724	conv5_3_0 (Conv2D)	(11, 11, 32)	4128
conv2_2_1 (Conv3D)	(24, 24, 29, 4)	724	conv5_4_0 (Conv2D)	(11, 11, 32)	4128
conv2_3_1 (Conv3D)	(24, 24, 29, 4)	724	conv5_1_1 (Conv2D)	(11, 11, 32)	9248
conv2_4_1 (Conv3D)	(24, 24, 29, 4)	724	conv5_2_1 (Conv2D)	(11, 11, 32)	9248
conv2_2 (Conv3D)	(23, 23, 14, 3)	544	conv5_3_1 (Conv2D)	(11, 11, 32)	9248
conv3_4_0 (Conv3D)	(23, 23, 14, 8)	264	conv5_4_1 (Conv2D)	(11, 11, 32)	9248
conv3_2_0 (Conv3D)	(23, 23, 14, 8)	264	conv5_5 (Conv2D)	(5, 5, 32)	4128
conv3_3_0 (Conv3D)	(23, 23, 14, 8)	264	flatten (Flatten)	(800)	0
conv3_1_0 (Conv3D)	(23, 23, 14, 8)	264	dense_1 (Dense)	(192)	153792
conv3_4_1 (Conv3D)	(23, 23, 14, 8)	1736	dropout_1 (Dropout)	(192)	0
conv3_2_1 (Conv3D)	(23, 23, 14, 8)	1736	dense_2 (Dense)	(128)	24704
conv3_3_1 (Conv3D)	(23, 23, 14, 8)	1736	dropout_2 (Dropout)	(128)	0
conv3_1_1 (Conv3D)	(23, 23, 14, 8)	1736	classifier (Dense)	(16)	2064
reshape (Reshape)	(22, 22, 224)	0	Total #Trainable params: 332,864		
conv3_3 (Conv2D)	(22, 22, 128)	28800			
conv4_1_0 (Conv2D)	(22, 22, 32)	4128			
conv4_2_0 (Conv2D)	(22, 22, 32)	4128			
conv4_3_0 (Conv2D)	(22, 22, 32)	4128			

Figure 5. Summary of the proposed model parameters

#### 4. RESULTS AND DISCUSSION

We conducted experiments to demonstrate how spatial information significantly influences HSI classification. Two HSI benchmark images, including the Indian Pines (IP) and the University of Pavia (PaviaU), were studied to evaluate the efficiency of our proposed model, and we compared our findings with various spectral-spatial HSI baseline approaches to determine the performance of the proposed model. Table 1 explains the significant features of each dataset used in our experiments, including the number of pixels, the number of spectral bands, wavelength range, spatial resolution, the number of classes, and the sensor, respectively. Further, to establish the efficacy of our model, we randomly chose  $N=5\%$  samples from each class to form the training data set for the IP and  $N=1\%$  the training sample size for the PaviaU dataset. Further, we designated the remaining reference samples as the testing data set. Here,  $x_i$  denotes the number of sample set sizes. Each experiment was repeated five times with randomly chosen training samples.

To quantitatively test the effectiveness of the proposed model and validity of the derived conclusions, three performance analysis metrics, i.e., overall accuracy (OA), average accuracy (AA), and kappa coefficient, were used. AA is the mean percentage of correctly classified pixels for each class. The kappa coefficient gives the percentage of correctly classified pixels about classification findings expected purely by chance.

Table 1. The hyperspectral dataset details [28]

HSI classification	#Pixels	Bands	Wavelength range (μm)	Spatial resolution	#Classes	Sensor
Indian Pines	145×145	200	0.4-2.5	20 m	16	AVIRIS
University of Pavia	610×340	103	0.43-0.86	1.3 m	9	ROSIS

To assess the achievement of our proposed approach, we compared the classification results of our model using the latest published baseline methods commonly used CNN-based HSI classification methods, including the 2D-CNN [29] (i.e., spatial approach), multi-scale 3D deep convolutional neural network (M3D-DCNN) [30], and residual hybrid spectral network (R-HybridSN) [31] (i.e., spectral-spatial methods). Tables 2 and 3 show the training samples for each dataset and the mean values, confidence interval, and classification performance for the two datasets. We can attest that our proposed model outperformed baseline methods from the achieved experimental results. Thus, we attribute this to the introduced DenseNet-based mechanism to model the semantic interdependencies in spatial and channel dimensions.

Figure 6 illustrates our model's OA against the baseline methods. Our proposed model has the best OA on the IP data set compared to M3D-DCNN and 2D-CNN methods but is slightly better than the

R-HybridSN method with a difference of 1.14% (OA =97.60%). Comparing the results of the PaviaU dataset (i.e., 5% train set) in contrast to the baseline methods, our model achieved excellent OA, i.e., 97.80%, as indicated in Table 3 and represented in Figure 5. Adopting the DenseNet layer on our model contributed highly to precise learning of the spatial-spectral classification of HSIs. Further, as evidenced in Tables 2 and 3, we notice most classes predicted with high accuracy. Table 4 lists the computational complexity results for various baseline methods. Because this image is roughly 145×145×200 pixels, it costs 26.9 seconds on the IP data set, outperforming the other compared methods.

Table 2. The classes detail information, the training and testing sample size, and the accuracy of different baseline methods, plus our proposed approach on the IP dataset

#	Name	#Samples	Train (5%)	Test (95%)	2D-CNN	M3D-DCNN	R-HybridSN	Ours
1	Alfalfa	46	2	44	7.95	27.5	45	73.81
2	Corn-Notill	1428	71	1357	70.69	59.15	95.45	94.14
3	Corn-Mintill	830	42	789	52.84	45.07	97.36	99.74
4	Corn	237	12	225	27.51	38.49	94.8	99.08
5	Grass-Pasture	483	24	459	90.44	70.33	98.85	97.3
6	Grass-Trees	730	37	694	98.59	97.2	99.32	99.66
7	Grass-Pasture-Mowed	28	1	27	10.37	18.52	95.56	96.31
8	Hay-Windrowed	478	24	454	99.96	98.04	100	100
9	Oats	20	1	19	16.32	25.79	65.26	88.89
10	Soybean-Notill	972	49	923	67.84	55.85	95.9	97.32
11	Soybean-Mintill	2455	123	2332	78.16	76.2	98.09	98.54
12	Soybean-Clean	593	30	563	42.01	33.89	89.15	95.24
13	Wheat	205	10	195	98.97	91.23	99.74	99.41
14	Woods	1265	63	1202	97.65	94.68	99.26	99.23
15	Buildings-Grass-Trees-Drives	386	19	367	62.62	42.37	87.66	98.03
16	Stone-Steel-Towers	93	5	88	76.02	49.32	88.18	91.76
AA (%)					62.37±1.64	57.73±6.52	90.60±1.53	95.15±0.31
OA (%)					75.47±1.64	68.88±3.77	96.46±0.33	97.60±0.34
Kappa (%)					0.718±0.01	0.642±0.045	0.960±0.004	0.973±0.56

Table 3. The classes detail information, the training and testing sample size, and the accuracy of different baseline methods, plus our proposed approach on the PaviaU dataset

#	Name	#Samples	Train (1%)	Test (99%)	2D-CNN	M3D-DCNN	R-HybridSN	Ours
1	Asphalt	6631	66	6565	96.88	90.56	96.94	97.4
2	Meadows	18649	186	18463	99.01	89.47	99.69	99.18
3	Gravel	2099	21	2078	75.08	59.11	87.17	88.34
4	Trees	3064	31	3033	87.74	93.25	89.15	89.88
5	Painted metal sheets	1345	13	1332	98.17	93.66	99.51	89.26
6	Bare Soil	5029	50	4979	75.51	69.63	98.44	98.54
7	Bitumen	1330	13	1317	61.32	65.71	95.82	98.25
8	Self-Blocking Bricks	3682	37	3645	80.61	78.35	93.28	93.51
9	Shadows	947	9	938	97.97	94.41	77.82	78.18
OA (%)					91.13±0.55	84.63±1.21	96.59±0.50	97.80±0.09
AA (%)					85.81±1.48	81.57±1.79	93.09±1.20	94.94±0.08
Kappa (%)					0.881±0.008	0.798±0.016	0.955±0.007	0.964±0.31

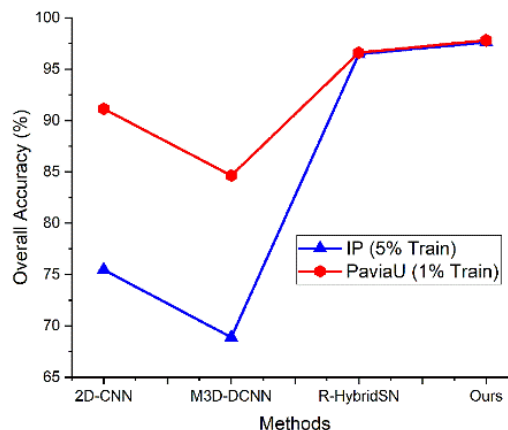


Figure 6. The OA for both IP (5% train set) and PaviaU (1% train set)



Table 4. The time (s) computation cost of various baseline methods in IP

Method	Training time (s)	Testing time (s)
2D-CNN	528.3	7.4
M3D-DCNN	660.6	126.3
R-HybridSN	464.3	60.9
Ours	430.8	26.9

## 5. CONCLUSION

In this study, we proposed a technique for learning high-level spectral-spatial features for HSI classification with insufficient labeled samples framework that jointly used dimension reduction and transfer learning techniques for learning the 3D spatial-spectral classification of HSIs. Our proposed framework fused a DenseNet transfer learning network for spectral-spatial feature learning, the PCA technique for extracting features from high-dimensional hyperspectral data sets, and the 3D and 2D convolutions. The experiments were conducted on two datasets, i.e., IP and the PaviaU. The results revealed our model's higher performance than other baseline methods that we compared our model. Thus, this method has good potential for precise HSI classification, often containing high dimensions of spectral-spatial features. With potential benefits for the environmental and agricultural domains, this suggested technique holds significant value in enhancing precision and efficacy in HSI categorization. We will include spatial information in the model, which was a limitation in the current work in the future to improve feature classification capability. Furthermore, because the value of spectral and spatial features differs depending on the material and scenario, it is also worth looking at how to automatically learn the proper weighting factor.

## REFERENCES




- [1] Q. Wang, Z. Meng, and X. Li, "Locality adaptive discriminant analysis for spectral-spatial classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 2077–2081, 2017, doi: 10.1109/LGRS.2017.2751559.
- [2] L. He, J. Li, A. Plaza, and Y. Li, "Discriminative low-rank gabor filtering for spectral-spatial hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1381–1395, 2017, doi: 10.1109/TGRS.2016.2623742.
- [3] P. Chen, L. Jiao, F. Liu, S. Gou, J. Zhao, and Z. Zhao, "Dimensionality reduction of hyperspectral imagery using sparse graph learning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 3, pp. 1165–1181, 2017, doi: 10.1109/JSTARS.2016.2606578.
- [4] A. Kianisarkaleh and H. Ghassemian, "Marginal discriminant analysis using support vectors for dimensionality reduction of hyperspectral data," *Remote Sensing Letters*, vol. 7, no. 12, pp. 1160–1169, 2016, doi: 10.1080/2150704X.2016.1222099.
- [5] F. Palsson, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Model-based fusion of multi-and hyperspectral images using PCA and wavelets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2652–2663, 2015, doi: 10.1109/TGRS.2014.2363477.
- [6] S. Shinde and H. Patidar, "Hyperspectral image classification using principle component analysis and deep convolutional neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 12, pp. 16491–16497, 2023, doi: 10.1007/s12652-022-03876-z.
- [7] G. Licciardi, P. R. Marpu, J. Chanussot, and J. A. Benediktsson, "Linear versus nonlinear PCA for the classification of hyperspectral data based on the extended morphological profiles," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 447–451, 2012, doi: 10.1109/LGRS.2011.2172185.
- [8] D. K. Pathak, S. K. Kalita, and D. K. Bhattacharya, "Hyperspectral image classification using support vector machine: a spectral spatial feature based approach," *Evolutionary Intelligence*, vol. 15, no. 3, pp. 1809–1823, 2022, doi: 10.1007/s12065-021-00591-0.
- [9] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013, doi: 10.1109/MGRS.2013.2244672.
- [10] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 45–54, 2014, doi: 10.1109/MSP.2013.2279179.
- [11] D. A. Landgrebe, *Signal theory methods in multispectral remote sensing*. Hoboken, USA: John Wiley & Sons, Inc., 2003, doi: 10.1002/0471723800.
- [12] A. Zare and K. C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 95–104, 2014, doi: 10.1109/MSP.2013.2279177.
- [13] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2013, doi: 10.1109/JPROC.2012.2197589.
- [14] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
- [15] M. Hamouda, K. S. Ettabaa, and M. S. Bouhlel, "Smart feature extraction and classification of hyperspectral images based on convolutional neural networks," *IET Image Processing*, vol. 14, no. 10, pp. 1999–2005, 2020, doi: 10.1049/iet-ipr.2019.1282.
- [16] L. Yang *et al.*, "FusionNet: A convolution-transformer fusion network for hyperspectral image classification," *Remote Sensing*, vol. 14, no. 16, 2022, doi: 10.3390/rs14164066.
- [17] B. Liu, K. Gao, A. Yu, L. Ding, C. Qiu, and J. Li, "ES2FL: Ensemble self-supervised feature learning for small sample classification of hyperspectral images," *Remote Sensing*, vol. 14, no. 17, 2022, doi: 10.3390/rs14174236.
- [18] X. Liu *et al.*, "H-RNet: Hybrid relation network for few-shot learning-based hyperspectral image classification," *Remote Sensing*, vol. 15, no. 10, 2023, doi: 10.3390/rs15102497.






- [19] S. Jia, S. Jiang, Z. Lin, N. Li, M. Xu, and S. Yu, "A survey: Deep learning for hyperspectral image classification with few labeled samples," *Neurocomputing*, vol. 448, pp. 179–204, 2021, doi: 10.1016/j.neucom.2021.03.035.
- [20] E. R. Kondal and S. S. Barpanda, "Hyperspectral image classification using Hyb-3D convolution neural network spectral partitioning," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 1, pp. 295–303, 2023, doi: 10.11591/ijeecs.v29.i1.pp295-303.
- [21] Z. Wu, Y. Gao, L. Li, J. Xue, and Y. Li, "Semantic segmentation of high-resolution remote sensing images using fully convolutional network with adaptive threshold," *Connection Science*, vol. 31, no. 2, pp. 169–184, 2019, doi: 10.1080/09540091.2018.1510902.
- [22] V. K. Shrivastava and M. K. Pradhan, "Hyperspectral remote sensing image classification using active learning," *Studies in Computational Intelligence*, vol. 907, pp. 133–152, 2021, doi: 10.1007/978-3-030-50641-4\_8.
- [23] Y. Guo, J. Ji, D. Shi, Q. Ye, and H. Xie, "Multi-view feature learning for VHR remote sensing image classification," *Multimedia Tools and Applications*, vol. 80, no. 15, pp. 23009–23021, 2021, doi: 10.1007/s11042-020-08713-z.
- [24] S. Pundir and J. A. Akshay, "EPM: Meta-learning method for remote sensing image classification," in *Machine Intelligence and Smart Systems*, Singapore: Springer, 2022, pp. 329–339, doi: 10.1007/978-981-16-9650-3\_25.
- [25] W. Ma *et al.*, "A novel adaptive hybrid fusion network for multiresolution remote sensing images classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022, doi: 10.1109/TGRS.2021.3062142.
- [26] N. Sathyanarayana and S. Singh, "Hybrid adaptive neural network for remote sensing image classification," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 2, pp. 2291–2300, 2024, doi: 10.11591/ijai.v13.i2.pp2291-2300.
- [27] S. Liu, S. Miao, J. Su, B. Li, W. Hu, and Y. D. Zhang, "UMAG-Net: A new unsupervised multiattention-guided network for hyperspectral and multispectral image fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7373–7385, 2021, doi: 10.1109/JSTARS.2021.3097178.
- [28] M. Graña, M. A. Veganzons, and B. Ayerdi, "Hyperspectral remote sensing scenes," *Group De Inteligencia Computacional*. 2011. Accessed: Jan. 06, 2020. [Online]. Available: [http://www.ehu.es/ccwintco/index.php?title=Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes)
- [29] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4959–4962, doi: 10.1109/IGARSS.2015.7326945.
- [30] M. He, B. Li, and H. Chen, "Multi-scale 3D deep convolutional neural network for hyperspectral image classification," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3904–3908, doi: 10.1109/ICIP.2017.8297014.
- [31] F. Feng, S. Wang, C. Wang, and J. Zhang, "Learning deep hierarchical spatial-spectral features for hyperspectral image classification based on residual 3D-2D CNN," *Sensors*, vol. 19, no. 23, 2019, doi: 10.3390/s19235276.

## BIOGRAPHIES OF AUTHORS



**Dr. Douglas Omwenga Nyabuga**    is a lecturer, postgraduate and undergraduate lecturer in the university level. He holds a Ph.D. in enterprise information systems and engineering from Donghua University, China, Masters in Computer Science from Periyar University, India. Published more than 15 academic papers in academic journals and conferences. He has been included in many papers by SCI and EI. His main research directions are; human-computer interaction (mainly education systems), image processing, remote sensing, pattern recognition, and machine learning. He is a distinguished reviewer at IET Image Processing Journal, IAES International Journal of Artificial Intelligence (IJ-AI), and a programme committee member for the international conference on artificial neural networks. His research interest focuses on utilizing computing and remote sensing techniques to solve problems related to natural resource management, ecology, natural disaster mapping, vegetation property extraction, and urban remote sensing. He can be contacted at email: [dnyabuga@mkurwanda.ac.rw](mailto:dnyabuga@mkurwanda.ac.rw).



**Godfrey Nyariki**    received the M.Sc. computer science degree from the Alagappa University, Karaikudi, India, with the dissertation "Incentivized peer - assisted streaming for on demand services". He is a lecturer of computer science in the Department of Information Technology, Mount Kigali University, Kigali, Rwanda since 2017. His research interests are in human-computer interaction, computer systems & networks, neural networks, pattern recognition, and image processing. He can be contacted at email: [gnyariki@mku.ac.ke](mailto:gnyariki@mku.ac.ke).